10/684,210 657

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００２年１２月１８日

出願番号
Application Number: 特願２００２－３６６３７４

[ST. 10/C]: [ＪＰ２００２－３６６３７４]

出願人
Applicant(s): 株式会社日立製作所

２００３年１０月 ８日

特許庁長官
Commissioner,
Japan Patent Office 今井康夫

出証番号 出証特２００３－３０８３０５７

【書類名】　特許願
【整理番号】　HI020826
【提出日】　平成14年12月18日
【あて先】　特許庁長官殿
【国際特許分類】　G06F　3/06
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番地２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　鈴木　進
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番地２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　長屋　賢紀
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番地２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　佐藤　孝夫
【特許出願人】
　【識別番号】　000005108
　【氏名又は名称】　株式会社日立製作所
【代理人】
　【識別番号】　100071283
　【弁理士】
　【氏名又は名称】　一色　健輔
【選任した代理人】
　【識別番号】　100084906
　【弁理士】
　【氏名又は名称】　原島　典孝

【選任した代理人】
　【識別番号】　100098523
　【弁理士】
　【氏名又は名称】　黒川　恵
【選任した代理人】
　【識別番号】　100112748
　【弁理士】
　【氏名又は名称】　吉田　浩二
【選任した代理人】
　【識別番号】　100110009
　【弁理士】
　【氏名又は名称】　青木　康
【手数料の表示】
　【予納台帳番号】　011785
　【納付金額】　21,000円
【提出物件の目録】
　【物件名】　明細書　1
　【物件名】　図面　1
　【物件名】　要約書　1
【プルーフの要否】　要

【書類名】　明細書

【発明の名称】　記憶デバイス制御装置の制御方法、記憶デバイス制御装置、及びプログラム

【特許請求の範囲】

【請求項１】　データを記憶するための記憶ボリュームを複数備える記憶デバイスと、前記データの入出力要求を行う情報処理装置とに接続され、前記情報処理装置から前記入出力要求を受信し、前記記憶ボリュームに対する前記データの入出力処理を行う記憶デバイス制御装置の制御方法であって、

ある前記記憶ボリュームと、前記ある記憶ボリュームに前記データが書き込まれると当該データの複製が書き込まれる他の前記記憶ボリュームとを対応付け、前記ある記憶ボリュームと前記他の記憶ボリュームとで構成されるペアを複数含んで構成されるペアグループを形成するステップと、

前記ペアグループを構成する各前記ペアの前記対応付けの解除を開始するステップと、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップと、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行うステップと

を備えることを特徴とする記憶デバイス制御装置の制御方法。

【請求項２】　前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行う前記ステップは、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記入出力要求が、前記対応付けが未解除の前記記憶ボリュームに対するものであるかどうかを判定し、前記対応付けが未解除の前記記憶ボリュームに対するものであった場合には、当該対応付けの解除を行ってから前記入出力処理を行うステップであること

を特徴とする請求項１に記載の記憶デバイス制御装置の制御方法。

【請求項３】　前記ペアグループを構成する前記各ペアの前記対応付けの解除を開始する前記ステップにおいて、前記記憶デバイス制御装置が備えるメモリに、当該ペアグループを構成する前記各ペアの前記対応付けの解除が開始されたことを示す識別情報を記憶しておき、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記ステップは、

前記識別情報を参照し、前記識別情報の内容に応じて、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップであること

を特徴とする請求項１に記載の記憶デバイス制御装置の制御方法。

【請求項４】　前記ペアグループを構成する前記各ペアの前記対応付けの解除を開始する前記ステップにおいて、前記記憶デバイス制御装置が備えるメモリに、当該ペアグループを構成する前記各ペアの前記対応付けの解除が開始される時刻を記憶しておき、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記ステップは、

前記メモリに記憶された時刻と、前記入出力要求に記述されている当該入出力要求が発行された時刻とを比較することにより、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップであること

を特徴とする請求項１に記載の記憶デバイス制御装置の制御方法。

【請求項５】　データを記憶するための記憶ボリュームを複数備える記憶デバイスと、前記データの入出力要求を行う情報処理装置とに接続され、前記情報処理装置から前記入出力要求を受信し、前記記憶ボリュームに対する前記データの入出力処理を行う記憶デバイス制御装置であって、

ある前記記憶ボリュームと、前記ある記憶ボリュームに前記データが書き込まれると当該データの複製が書き込まれる他の前記記憶ボリュームとを対応付け、前記ある記憶ボリュームと前記他の記憶ボリュームとで構成されるペアを複数含んで構成されるペアグループを形成する手段と、

前記ペアグループを構成する各前記ペアの前記対応付けの解除を開始する手段と、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する手段と、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行う手段と

を備えることを特徴とする記憶デバイス制御装置。

【請求項６】　前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行う前記手段は、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記入出力要求が、前記対応付けが未解除の前記記憶ボリュームに対するものであるかどうかを判定し、前記対応付けが未解除の前記記憶ボリュームに対するものであった場合には、当該対応付けの解除を行ってから前記入出力処理を行う手段であること

を特徴とする請求項５に記載の記憶デバイス制御装置。

【請求項７】　メモリと、

前記ペアグループを構成する前記各ペアの前記対応付けの解除が開始されたことを示す識別情報を前記メモリに記憶する手段とをさらに備え、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記手段は、

前記識別情報を参照し、前記識別情報の内容に応じて、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する手段であること

を特徴とする請求項５に記載の記憶デバイス制御装置。

【請求項８】　メモリと、

前記ペアグループを構成する前記各ペアの前記対応付けの解除が開始される時刻を前記メモリに記憶する手段とをさらに備え、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記手段は、

前記メモリに記憶された時刻と、前記入出力要求に記述されている当該入出力要求が発行された時刻とを比較することにより、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する手段であること

を特徴とする請求項５に記載の記憶デバイス制御装置。

【請求項９】　データを記憶するための記憶ボリュームを複数備える記憶デバイスと、前記データの入出力要求を行う情報処理装置とに接続され、前記情報処理装置から前記入出力要求を受信し、前記記憶ボリュームに対する前記データの入出力処理を行う記憶デバイス制御装置に、

ある前記記憶ボリュームと、前記ある記憶ボリュームに前記データが書き込まれると当該データの複製が書き込まれる他の前記記憶ボリュームとを対応付け、前記ある記憶ボリュームと前記他の記憶ボリュームとで構成されるペアを複数含んで構成されるペアグループを形成するステップと、

前記ペアグループを構成する各前記ペアの前記対応付けの解除を開始するステップと、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップと、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行うステップと

を実行させるためのプログラム。

【請求項１０】　前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行う前記ステップは、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記入出力要求が、前記対応付けが未解除の前記記憶ボリュームに対するものであるかどうかを判定し、前記対応付けが未解

除の前記記憶ボリュームに対するものであった場合には、当該対応付けの解除を行ってから前記入出力処理を行うステップであること
を特徴とする請求項９に記載のプログラム。

【請求項１１】　前記ペアグループを構成する前記各ペアの前記対応付けの解除を開始する前記ステップにおいて、前記記憶デバイス制御装置が備えるメモリに、当該ペアグループを構成する前記各ペアの前記対応付けの解除が開始されたことを示す識別情報を記憶しておき、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記ステップは、

前記識別情報を参照し、前記識別情報の内容に応じて、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップであること
を特徴とする請求項９に記載のプログラム。

【請求項１２】　前記ペアグループを構成する前記各ペアの前記対応付けの解除を開始する前記ステップにおいて、前記記憶デバイス制御装置が備えるメモリに、当該ペアグループを構成する前記各ペアの前記対応付けの解除が開始される時刻を記憶しておき、

前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定する前記ステップは、

前記メモリに記憶された時刻と、前記入出力要求に記述されている当該入出力要求が発行された時刻とを比較することにより、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップであること
を特徴とする請求項９に記載のプログラム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、記憶デバイス制御装置の制御方法、記憶デバイス制御装置、及びプログラムに関する。

【０００２】

【従来の技術】

通信可能に接続された情報処理装置とディスクアレイ装置とを含む記憶装置システムにおいて、データの複製元となるプライマリボリューム（正ボリューム）と複製先となるセカンダリボリューム（副ボリューム）とでペアを形成し、プライマリボリュームからセカンダリボリュームへのリアルタイムなデータの複製により、プライマリボリュームのデータを二重に管理する複製管理機能が知られている。

【０００３】

【特許文献１】

特開２００１－３１８８３３号公報

【０００４】

【発明が解決しようとする課題】

しかしながら、例えば情報処理装置とディスクアレイ装置との間で授受されるデータが１つのプライマリボリュームに収まりきらない場合等、複数のプライマリボリュームに跨ってデータが格納される場合がある。このような場合に例えばデータのバックアップをとろうとすると、複数のペアに対してペア状態を解除する必要がある。ここで、個々のペアに対するペア状態の解除を行っている間にプライマリボリュームに対するデータの更新があると、ペア状態解除済みのプライマリボリュームに対するデータの更新はセカンダリボリュームに反映されないのに対し、ペア状態が未解除のプライマリボリュームに対するデータの更新はセカンダリボリュームに反映される場合がある。

【０００５】

そこで本発明は、複数の記憶ボリュームに格納されたデータの整合性を保ちつつデータの複製管理を行うことができる記憶デバイス制御装置の制御方法、記憶デバイス制御装置、及びプログラムを提供することを主たる目的とする。

【０００６】

【課題を解決するための手段】

データを記憶するための記憶ボリュームを複数備える記憶デバイスと、前記デ

ータの入出力要求を行う情報処理装置とに接続され、前記情報処理装置から前記入出力要求を受信し、前記記憶ボリュームに対する前記データの入出力処理を行う記憶デバイス制御装置の制御方法であって、ある前記記憶ボリュームと、前記ある記憶ボリュームに前記データが書き込まれると当該データの複製が書き込まれる他の前記記憶ボリュームとを対応付け、前記ある記憶ボリュームと前記他の記憶ボリュームとで構成されるペアを複数含んで構成されるペアグループを形成するステップと、前記ペアグループを構成する各前記ペアの前記対応付けの解除を開始するステップと、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものであるかどうかを判定するステップと、前記情報処理装置からの前記入出力要求が、前記対応付けの解除を開始した後に発行されたものである場合には、前記対応付けの解除を行ってから前記入出力処理を行うステップとを備える。

【０００７】

ここで、記憶デバイスと記憶デバイス制御装置とは、上述したディスクアレイ装置を構成する。そして情報処理装置とディスクアレイ装置とは、上述した記憶装置システムを構成する。

【０００８】

記憶ボリュームとは、ディスクアレイ装置又は記憶デバイスが備えるディスクドライブにより提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む、ディスクアレイ装置又は記憶デバイスにより提供される記憶リソースである。

【０００９】

また、上記ある記憶ボリュームと他の記憶ボリュームとが対応付けられている状態が上述した「ペア状態」である。

【００１０】

【発明の実施の形態】

以下、本発明の実施の形態について図面を用いて詳細に説明する。

＜第１の実施の形態＞

＝＝＝全体構成例＝＝＝

まず、本実施の形態に係る記憶装置システムの全体構成を示すブロック図を図１に示す。

【００１１】

情報処理装置１００はＣＰＵ（Central Processing Unit）やメモリを備えたコンピュータである。情報処理装置１００が備えるＣＰＵにより各種プログラムが実行されることにより、様々な機能が実現される。情報処理装置１００は例えば銀行の自動預金預け払いシステムや航空機の座席予約システム等における中枢コンピュータとして利用される。

【００１２】

また情報処理装置１００は記憶デバイス制御装置２００と接続されており、記憶デバイス制御装置２００との間で通信を行う。記憶デバイス制御装置２００に対してデータ入出力コマンド（データの入出力要求）を送信することにより、記憶デバイス３００に記憶されているデータの読み出し、あるいは記憶デバイス３００に対するデータの書き込みを行う。また情報処理装置１００は、記憶デバイス制御装置２００との間で記憶デバイス３００を管理するための各種コマンドの授受も行う。例えば記憶デバイス３００が備える記憶ボリュームに記憶されるデータの複製管理のためのコマンド等である。

【００１３】

情報処理装置１００の構成を示すブロック図を図２に示す。

情報処理装置１００は、ＣＰＵ１１０、メモリ１２０、ポート１３０、記録媒体読取装置１４０、入力装置１５０、出力装置１６０を備える。

【００１４】

ＣＰＵ１１０は情報処理装置１００の全体の制御を司るもので、メモリ１２０に格納されたプログラムを実行することにより各種機能を実現する。記録媒体読取装置１４０は、記録媒体１７０に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータはメモリ１２０に格納される。従って、例えば記録媒体１７０に記録された記憶デバイス管理プログラム１２１やアプリケーションプログラム１２２を、記録媒体読取装置１４０を用いて上記記録媒体１７０から読み取って、メモリ１２０に格納するようにすること

ができる。記録媒体１７０としてはフレキシブルディスクやＣＤ－ＲＯＭ、半導体メモリ等を用いることができる。記録媒体読取装置１４０は情報処理装置１００に内蔵されている形態とすることもできるし、外付されている形態とすることもできる。入力装置１５０はオペレータ等による情報処理装置１００へのデータ入力等のために用いられる。入力装置１５０としては例えばキーボードやマウス等が用いられる。出力装置１６０は情報を外部に出力するための装置である。出力装置１６０としては例えばディスプレイやプリンタ等が用いられる。ポート１３０は記憶デバイス制御装置２００と通信を行うための装置である。また他の情報処理装置１００との間で通信を行うために使用することもできる。この場合、例えば記憶デバイス管理プログラム１２１やアプリケーションプログラム１２２を、ポート１３０を介して他の情報処理装置１００から受信して、メモリ１２０に格納するようにすることもできる。

【００１５】

記憶デバイス管理プログラム１２１は、記憶デバイス３００が備える記憶ボリュームに記憶されるデータの複製管理を行うためのプログラムである。複製管理を行うための各種コマンドを記憶デバイス制御装置２００に対して送信することにより複製管理が行われる。

【００１６】

アプリケーションプログラム１２２は、情報処理装置１００により各種機能を実現するためのプログラムである。例えば上述した銀行の自動預金預け払いシステムを実現するための機能や航空機の座席予約システムを実現するための機能等である。

【００１７】

次に図１に戻って記憶デバイス制御装置２００について説明する。記憶デバイス制御装置２００は、情報処理装置１００から受信したコマンドに従って記憶デバイス３００に対する制御を行う。例えば情報処理装置１００からデータの入出力要求を受信して、記憶デバイス３００が備える記憶ボリュームに記憶されているデータの入出力処理を行う。

【００１８】

記憶デバイス制御装置２００は、チャネルアダプタ２１０、キャッシュメモリ２２０、共有メモリ２３０、ディスクアダプタ２４０、管理端末（ＳＶＰ：Service Processor）２６０、接続部２５０を備える。

【００１９】

チャネルアダプタ２１０は情報処理装置１００との通信インタフェースを備え、情報処理装置１００との間でデータ入出力コマンド等を授受する。

【００２０】

チャネルアダプタ２１０の構成を示すブロック図を図３に示す。

チャネルアダプタ２１０は、ＣＰＵ２１１、キャッシュメモリ２１２、制御メモリ２１３、ポート２１５、バス２１６を備えている。

【００２１】

ＣＰＵ２１１はチャネルアダプタ２１０の全体の制御を司るものであり、制御メモリ２１３に格納された制御プログラム２１４を実行する。制御メモリ２１３に格納された制御プログラム２１４が実行されることにより本実施の形態に係るデータの複製管理が実現される。キャッシュメモリ２１２は情報処理装置１００との間で授受されるデータやコマンド等を一時的に格納するためのメモリである。ポート２１５は情報処理装置１００との間の通信や記憶デバイス制御装置２００内部の他の装置との間の通信を行うための通信インタフェースである。バス２１６はこれらの装置を相互に接続する。

【００２２】

再び図１に戻る。キャッシュメモリ２２０は、チャネルアダプタ２１０とディスクアダプタ２４０との間で授受されるデータを一時的に格納するための記憶メモリである。すなわち、例えばチャネルアダプタ２１０が情報処理装置１００から受信したデータ入出力コマンドが書き込みコマンドであった場合には、チャネルアダプタ２１０は当該コマンドを共有メモリ２３０に書き込むと共に、情報処理装置１００から受信した書き込みデータをキャッシュメモリ２２０に書き込む。そうすると、ディスクアダプタ２４０は共有メモリ２３０に書き込まれた当該書き込みコマンドに従って、キャッシュメモリ２２０から書き込みデータを読み出して記憶デバイス３００に書き込む。

【００２３】

管理端末２６０は記憶デバイス制御装置２００や記憶デバイス３００を保守・管理するための情報処理装置の一形態である。例えばチャネルアダプタ２１０において実行される制御プログラム２１４の変更は管理端末２６０を用いて行われる。管理端末２６０は記憶デバイス制御装置２００に内蔵されている形態とすることもできるし、別体とすることもできる。また管理端末２６０は、記憶デバイス制御装置２００及び記憶デバイス３００の保守・管理を専用に行う情報処理装置とすることもできるし、汎用の情報処理装置に保守・管理機能を持たせたものとすることもできる。管理端末２６０の構成は図２に示した情報処理装置１００の構成と同様である。すなわち管理端末２６０は、ＣＰＵ１１０やメモリ１２０、ポート１３０、記録媒体読取装置１４０、入力装置１５０、出力装置１６０を備えている。従って、チャネルアダプタ２１０において実行される制御プログラム２１４は、管理端末２６０の記録媒体読取装置１４０を用いて記録媒体１７０から読み取るようにすることもできるし、管理端末２６０のポート１３０を介して接続される他の情報処理装置１００から受信するようにすることもできる。

【００２４】

ディスクアダプタ２４０はチャネルアダプタ２１０からの指示に従い、記憶デバイス３００の制御を行う。

【００２５】

記憶デバイス３００は、情報処理装置１００に提供するための記憶ボリュームを備えている。記憶ボリュームとは、記憶デバイス３００が備えるディスクドライブにより提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む、記憶デバイス３００により提供される記憶リソースである。ディスクドライブとしては、例えばハードディスク装置やフレキシブルディスク装置、半導体記憶装置等様々なものを採用することができる。またディスクアダプタ２４０と記憶デバイス３００との間は図１のように直接に接続される形態とすることもできるし、ネットワークを介して接続するようにすることもできる。さらに記憶デバイス３００は記憶デバイス制御装置２００と一体型として構成されることもできる。

【００２６】

共有メモリ２３０はチャネルアダプタ２１０とディスクアダプタ２４０の両方からアクセスが可能である。データ入出力要求コマンドの受け渡しに利用される他、記憶デバイス制御装置２００や記憶デバイス３００の管理情報等が記憶される。本実施の形態においては、図４に示すようにコンシステンシグループ管理テーブル２３１、ペア管理テーブル２３２が共有メモリ２３０に記憶される。

【００２７】

＝＝＝ペア管理テーブル＝＝＝

ペア管理テーブル２３２は、記憶デバイス３００に記憶されるデータの複製管理を行うためのテーブルである。ペア管理テーブル２３２は、「ペア」欄、「正ボリューム」欄、「副ボリューム」欄、「ペア状態」欄、「コンシステンシグループ」欄を備える。

【００２８】

「ペア」欄はペアの名称を示す。ペアとは２つの記憶ボリュームにより形成される記憶ボリュームの組み合わせをいう。ペアを形成している記憶ボリュームの例を図５に示す。図５においては２組のペア、すなわちペアＡとペアＢとが示されている。ペアを形成する記憶ボリュームは、一方をプライマリボリューム、他方をセカンダリボリュームとして管理される。図５においてはプライマリボリュームを「正」、セカンダリボリュームを「副」と記されている。なお、一つのプライマリボリュームに対して複数のセカンダリボリュームを組み合わせることも可能である。

【００２９】

図４のペア管理テーブル２３２に戻ると、「正ボリューム」欄は当該ペアのプライマリボリュームを示す。「副ボリューム」欄は当該ペアのセカンダリボリュームを示す。

【００３０】

「ペア状態」欄は当該ペアの状態を示す。ペアの状態としては、「ペア中」、「スプリット中」、「リシンク中」がある。

「ペア中」の場合は、情報処理装置１００からプライマリボリュームに書き込

まれたデータがセカンダリボリュームにも反映される。このようなプライマリボリュームとセカンダリボリュームとの対応付けにより、プライマリボリュームに記憶されている内容とセカンダリボリュームに記憶されている内容との同一性を確保することができる。

【００３１】

「スプリット中」の場合は、情報処理装置１００からプライマリボリュームにデータが書き込まれてもセカンダリボリュームには反映されない。すなわち「スプリット中」の状態は、「ペア中」におけるプライマリボリュームとセカンダリボリュームとの対応付けが解除された状態である。従ってプライマリボリュームの内容とセカンダリボリュームの内容の同一性は保証されない。しかし「スプリット中」はセカンダリボリュームの内容が更新されないので、この間にデータのバックアップをとることが可能となる。例えば磁気テープ等にセカンダリボリュームに記憶されているデータを保存するようにする。これにより、情報処理装置１００がそれまで行っていた業務をプライマリボリュームにより継続したまま、データのバックアップをとることが可能となる。

【００３２】

「リシンク中」は「スプリット中」から「ペア中」に移行する途中の状態である。すなわち、「スプリット中」にプライマリボリュームに対してなされたデータの更新をセカンダリボリュームに反映している状態である。反映が完了すると当該ペアの状態は「ペア中」になる。

【００３３】

上記ペアの形成や、「ペア中」の記憶ボリュームの「スプリット中」への移行、あるいは「スプリット中」の記憶ボリュームの「ペア中」への移行は、記憶デバイス管理プログラム１２１が実行されている情報処理装置１００に対して入力装置１５０からオペレータが指示入力を与えることにより行われる。オペレータにより与えられた指示入力は記憶デバイス制御装置２００のチャネルアダプタ２１０に送信される。チャネルアダプタ２１０は制御プログラム２１４を実行して、上記指示に従って、ペアの形成やペア状態の変更を行う。そして形成したペアのペア状態に応じて、例えば「ペア中」のペアに対してはプライマリボリューム

の更新データの複製をセカンダリボリュームへ反映する等、記憶ボリュームに対する制御を行う。

【００３４】

チャネルアダプタ２１０によるペア状態の変更は、各ペアに対して逐次に行われる。なぜならば、例えば上述したように一つのプライマリボリュームに対して複数のセカンダリボリュームを組み合わせてペアを形成することもできるので、かかる場合に同時に複数のペア状態が変化すると、プライマリボリュームの管理が複雑になるからである。

【００３５】

なお、上記ペアの形成やペア状態の移行は、情報処理装置１００がオペレータから受けた指示により開始される他、オペレータからの指示がなくても、例えば所定の時刻になると自動的にペア状態の移行が行われるようにすることもできる。またポート１３０を介して接続される他の情報処理装置１００から指示を受信することにより行われるようにすることもできる。

【００３６】

＝＝＝コンシステンシグループ＝＝＝

「コンシステンシグループ」欄には、当該ペアが属するコンシステンシグループ（ペアグループ）の番号が記載される。ここでコンシステンシグループとは、スプリット状態への移行が同時に行われるように制御される複数の記憶ボリュームのペアにより構成されるグループをいう。つまり、上述のように複数のペアに対するペア状態の変更は各ペアに対して逐次に行われるが、同一コンシステンシグループに属する複数のペアに対しては、スプリット状態への移行が同時に行われるように制御される（以下、スプリット状態への移行の同時性という）。

【００３７】

例えば、同一コンシステンシグループに属する複数のペアを順次ペア状態からスプリット状態に変更している間に情報処理装置１００からデータの書き込みがあった場合を考える。仮にコンシステンシグループが形成されていなければ、当該データの書き込みがスプリット状態への移行が完了したペアのプライマリボリュームに対するものであった場合には当該書き込みデータはセカンダリボリュー

ムには反映されないが、スプリット状態への移行が完了していないペアのプライマリボリュームに対するものであった場合には当該書き込みデータはセカンダリボリュームに反映されてしまう。しかしながら、コンシステンシグループが形成されている場合には、スプリット状態への移行が完了している、していないに拘わらず、当該書き込みデータはセカンダリボリュームへは反映されない。なぜならば、当該データの書き込みは、コンシステンシグループのペアに対するスプリットを開始（対応付けの解除を開始）した後に発行されたものだからである。

【００３８】

このように複数のペアに対してコンシステンシグループを形成することは、特に、例えば書き込みデータが一つの記憶ボリュームに収まりきらないほど大きなものある場合や、一つのファイルデータを複数の記憶ボリュームに格納するように制御している場合等、一つのデータが複数の記憶ボリュームに跨って格納されている場合に有効である。

【００３９】

なお、コンシステンシグループによる各ペアのスプリット状態への移行の同時性の確保は、情報処理装置１００からセカンダリボリュームへのデータの書き込み要求又は読み出し要求があった場合にも有効である。

すなわち、仮にコンシステンシグループが形成されていなければ、スプリット状態への移行が完了したペアのセカンダリボリュームに対してはデータの書き込み又は読み出しが可能であるのに対し、スプリット状態への移行が完了していないペアのセカンダリボリュームに対してはデータの書き込み又は読み出しが禁止されているからである。

【００４０】

本実施の形態においては、上記コンシステンシグループによる各ペアのスプリット状態への移行の同時性の確保は、コンシステンシグループ管理テーブル２３１の一括スプリット受領フラグ（識別情報）を用いることにより行われる。図６に示すフローチャートを用いて処理の流れを説明する。

【００４１】

＝＝＝処理の流れ＝＝＝

なお以下の処理は、本実施の形態に係る各種の動作を行うためのコードから構成される制御プログラム２１４（プログラム）を、チャネルアダプタ２１０が備えるＣＰＵ２１１が実行することにより実現される。

【００４２】

まず、チャネルアダプタ２１０は情報処理装置１００から、あるコンシステンシグループに対するペア分割要求（スプリット命令）を受領する（S1000）。そうするとチャネルアダプタ２１０は、共有メモリ２３０のコンシステンシグループ管理テーブル２３１の一括スプリット受領フラグを”ＯＮ”にする（S1001）。次に、当該コンシステンシグループに属する未分割のペアのスプリット状態への移行を開始する（S1003）。具体的には、当該ペアのプライマリボリュームとセカンダリボリュームとの対応付けを解除し、プライマリボリュームへの書き込みデータのセカンダリボリュームへの反映を中止する。そしてペア管理テーブル２３２の当該ペアのペア状態欄を「スプリット中」に変更する（S1004）。以上の処理を同一コンシステンシグループ内の全てのペアに対して繰り返す。同一コンシステンシグループ内の全てのペアが「スプリット中」になったら（S1005）、一括スプリットフラグを”ＯＦＦ”にして、処理を終了する。

【００４３】

上記処理の途中で情報処理装置１００からデータのリードライト要求を受領した場合には、当該データのリードライト要求がペア未分割の記憶ボリューム、すなわちペア状態が「ペア中」の記憶ボリューム（対応付けが未解除の記憶ボリューム）に対するものであるかどうかをチェック（判定）する（S1006）。「ペア中」の記憶ボリュームに対するものである場合には、まず当該記憶ボリュームのペアを「スプリット中」に移行させる（S1007）。そしてペア管理テーブル２３２の当該ペアのペア状態欄を「スプリット中」に変更し（S1008）、その後データのリードライト処理（入出力処理）を実行する（S1009）。

【００４４】

一方（S1006）において、当該データのリードライト要求がペア分割済み、すなわちペア状態が「スプリット中」の記憶ボリュームに対するものであった場合には、当該記憶ボリュームに対するリードライト処理を実行する（S1009）。

【００４５】

このようにすることにより、同一コンシステンシグループ内の各ペアのスプリット状態への移行の同時性を確保することが可能となる。

【００４６】

なお、図６に示したフローチャートにおいては、コンシステンシグループ内のペアを順次分割している途中で情報処理装置１００からデータのリードライト要求を受領した場合には、当該データのリードライト要求がペア未分割の記憶ボリュームに対するものであるかどうかをチェックし（S1006）リードライト処理を実行（S1009）しているが、コンシステンシグループ内のペアを順次分割している途中の状態で情報処理装置１００からデータのリードライト要求を受領しても、リードライト処理を実行しないようにすることもできる。この場合には、コンシステンシグループ内の全てのペアの分割が完了し一括スプリットフラグが”ＯＦＦ”にされた後に、当該リードライト処理が実行される。

【００４７】

以上の処理を詳細に示したフローチャートを図７に示す。

まず、情報処理装置１００から送信されたコマンドによりペアＡ、ペアＢに対してコンシステンシグループが形成される（S2000乃至S2002）。コマンドは例えばオペレータにより情報処理装置１００の入力装置１５０から入力される。情報処理装置１００に入力されたコマンドは、記憶デバイス管理プログラム１２１によりチャネルアダプタ２１０に送信される。図７に示す”paircreate -g GRP0”はコマンドの一例である。このコマンドを受信することによりチャネルアダプタ２１０はコンシステンシグループ０を形成する。また、チャネルアダプタ２１０は共有メモリ２３０のペア管理テーブル２３２、及びコンシステンシグループ管理テーブル２３１に所定のデータを記録する。その様子が図４に示される。ただし、図４においてはペア管理テーブル２３２のペアＡのペア状態欄は「スプリット中」と記載されているが、この段階ではペアＡのペア状態の欄は「ペア中」である。またコンシステンシグループ管理テーブル２３１のコンシステンシグループ０の一括スプリット受領フラグ欄が”ＯＮ”と記載されているが、この段階では”ＯＦＦ”である。

【００４８】

この状態で情報処理装置１００から、ペアＡを形成する記憶ボリューム１に対するデータのリードライト要求（R/W１）を受領した場合には（S2008）、チャネルアダプタ２１０は、通常の通りにリードライト処理を実行する（S2009）。コンシステンシグループ管理テーブル２３１のコンシステンシグループ０に対する一括スプリット受領フラグ欄が”ＯＦＦ”だからである。

【００４９】

その後、情報処理装置１００からのコマンドによりペアＡ、ペアＢにより構成されるコンシステンシグループ０に対してペアの分割が指示される（S2003）。図７に示す”pairsplit -g GRP0”はコマンドの一例である。このコマンドも、例えばオペレータにより情報処理装置１００の入力装置１５０から入力される。

【００５０】

そうすると、チャネルアダプタ２１０は、共有メモリ２３０のコンシステンシグループ管理テーブル２３１のコンシステンシグループ０に対する一括スプリット受領フラグを”ＯＮ”にして（S2004）、各ペアに対するスプリットを順次開始する（S2005、S2006）。ペアＡに対するスプリットが完了した状態が図４のペア管理テーブル２３２に示されている。全てのペアに対するスプリットが完了したら、一括スプリット受領フラグを”ＯＦＦ”にして処理を終了する（S2007）。

【００５１】

ここで、コンシステンシグループ０に対するペア分割のコマンドを情報処理装置１００から受領した後（S2003）であって、一括スプリット受領フラグを”ＯＮ”にする（S2004）までの間に、情報処理装置１００からペアＢを形成する記憶ボリューム３に対するデータのリードライト要求（R/W２）を受領した場合には（S2010）、チャネルアダプタ２１０は、通常通りに当該リードライト処理を実行する（S2011）。コンシステンシグループ管理テーブル２３１のコンシステンシグループ０に対する一括スプリット受領フラグ欄が”ＯＦＦ”だからである。

【００５２】

しかし、一括スプリット受領フラグを”ＯＮ”した後に（S2004）情報処理装置１００からペアＢを形成する記憶ボリューム３に対するデータのリードライト要

求（R/W３）を受領した場合には（S2012）、まずペアＢのスプリットを実行し（S2013）、その後リードライト処理を実行する（S2014）。

【００５３】

このようにチャネルアダプタ２１０は、情報処理装置１００からリードライト要求を受領した際に一括スプリット受領フラグを参照することにより、当該リードライト要求が当該コンシステンシグループを構成する各ペアのペア状態の解除を開始した後に発行されたものであるかどうかを判定する。

【００５４】

（S2005）においてペアＡに対するスプリットが完了した後に、情報処理装置１００からペアＡを形成する記憶ボリューム１に対するデータのリードライト要求（R/W４）を受領した場合には（S2015）、チャネルアダプタ２１０は当該リードライト処理を実行する（S2016）。ペア管理テーブル２３２のペアＡに対するペア状態欄が「スプリット中」となっているので、ペアＡのスプリットが完了していることが分かるからである。

【００５５】

なお、（S2013）によりデータのリードライト処理の際にペアＢのスプリットが完了しているので、（S2005）においてはペアＢのスプリット処理は行われない。

【００５６】

このように本実施の形態においては、一括スプリット受領フラグを設けることにより、同一コンシステンシグループ内の各ペアのスプリット状態への移行の同時性を確保することができる。

【００５７】

＜第２の実施の形態＞

＝＝＝コンシステンシグループ管理テーブル＝＝＝

次に、第２の実施の形態について説明する。

第１の実施の形態との相違点は、図８に示すようにコンシステンシグループ管理テーブル２３１においてスプリット開始時刻が記録される点である。図８の例では、コンシステンシグループ０に属するペアは、時刻１２時になるとスプリッ

トが開始されることが示されている。そして、コンシステンシグループ０に属する全てのペアのスプリットが完了すると、スプリット開始時刻欄は”－”に変更される。

【００５８】

スプリットの開始時刻は、情報処理装置１００から送信されるコマンドにより指示されるが、具体的な時刻が指示されるのではなく、直ちにスプリットを開始せよとの指示を行えるようにすることも可能である。この場合には、スプリット開始時刻欄には現在の時刻が記録される。

【００５９】

本実施の形態においては、チャネルアダプタ２１０は情報処理装置１００からデータのリードライトコマンドを受領すると、当該リードライトコマンドに記録されているリードライトコマンドの発行時刻と、コンシステンシグループ管理テーブル２３１のスプリット開始時刻欄に記載された時刻とを比較する。そしてリードライトコマンドの発行時刻の方が遅い場合には、当該リードライト処理はスプリット実行後に行われる。

これにより、同一コンシステンシグループ内の各ペアのスプリット状態への移行の同時性を確保することができる。

【００６０】

＝＝＝処理の流れ＝＝＝

具体的な処理の流れを図９のフローチャートを用いて説明する。

なお以下の処理は、本実施の形態に係る各種の動作を行うためのコードから構成される制御プログラム２１４（プログラム）を、チャネルアダプタ２１０が備えるＣＰＵ２１１が実行することにより実現される。

【００６１】

まず、チャネルアダプタ２１０は情報処理装置１００から、あるコンシステンシグループに対するペア分割要求（スプリット命令）を受領する（S3000）。そうするとチャネルアダプタ２１０は、当該スプリット命令に記録されているスプリット開始時刻を共有メモリ２３０のコンシステンシグループ管理テーブル２３１のスプリット開始時刻欄に記録する（S3001）。次に、現在の時刻とスプリッ

ト開始時刻とを比較し、スプリット開始時刻を経過していないかどうかをチェックする（S3003）。スプリット開始時刻を経過している場合には、当該コンシステンシグループに属する未分割のペアのスプリット状態への移行を開始する（S3004）。具体的には、当該ペアのプライマリボリュームとセカンダリボリュームとの対応付けを解除し、プライマリボリュームへの書き込みデータのセカンダリボリュームへの反映を中止する。そしてペア管理テーブル２３２の当該ペアのペア状態欄を「スプリット中」に変更する（S3005）。以上の処理を同一コンシステンシグループ内の全てのペアに対して繰り返す。同一コンシステンシグループ内の全てのペアが「スプリット中」になったら（S3006）、スプリット開始時刻欄を”－”に変更して処理を終了する（S3007）。

【００６２】

上記処理の途中で情報処理装置１００からデータのリードライト要求を受領した場合には、当該データのリードライト要求がペア未分割の記憶ボリューム、すなわちペア状態が「ペア中」の記憶ボリューム（対応付けが未解除の記憶ボリューム）に対するものであるかどうかをチェック（判定）する（S3008）。「ペア中」の記憶ボリュームに対するものである場合には、当該リードライトコマンドに記録されているリードライトコマンドの発行時刻と、スプリット開始時刻とを比較する（S3010）。そしてリードライトコマンドの発行時刻の方が遅い場合には、まず当該記憶ボリュームのペアを「スプリット中」に移行させる（S3011）。そしてペア管理テーブル２３２の当該ペアのペア状態欄を「スプリット中」に変更し（S3012）、その後データのリードライト処理（入出力処理）を実行する（S3013）。

【００６３】

一方（S3008）において当該データのリードライト要求がペア分割済み、すなわちペア状態が「スプリット中」の記憶ボリュームに対するものであった場合、または（S3010）において当該リードライト要求に記録されているリードライトコマンドの発行時刻がスプリット開始時刻よりも早い場合には、当該記憶ボリュームに対するリードライト処理を実行する（S3009）。

【００６４】

このようにすることにより、同一コンシステンシグループ内の各ペアのスプリット状態への移行の同時性を確保することが可能となる。

【００６５】

なお、図９に示したフローチャートにおいては、コンシステンシグループ内のペアを順次分割している途中で情報処理装置１００からデータのリードライト要求を受領した場合には、当該データのリードライト要求がペア未分割の記憶ボリュームに対するものであるかどうかをチェックし（S3008）リードライト処理を実行（S3009、S3013）　しているが、コンシステンシグループ内のペアを順次分割している途中の状態で情報処理装置１００からデータのリードライト要求を受領しても、リードライト処理を実行しないようにすることもできる。この場合には、コンシステンシグループ内の全てのペアの分割が完了しスプリット開始時刻欄が”－”に変更された後に、当該リードライト処理が実行される。

【００６６】

以上本実施の形態について説明したが、上記実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明はその趣旨を逸脱することなく変更、改良され得ると共に、本発明にはその等価物も含まれる。

【００６７】

【発明の効果】

記憶デバイス制御装置の制御方法、記憶デバイス制御装置、及びプログラムを提供することができる。

【図面の簡単な説明】

【図１】　本実施の形態に係る記憶装置システムの全体構成を示すブロック図である。

【図２】　本実施の形態に係る情報処理装置の構成を示すブロック図である。

【図３】　本実施の形態に係る記憶デバイス制御装置におけるチャネルアダプタの構成を示すブロック図である。

【図４】　本実施の形態に係る記憶デバイス制御装置における共有メモリに記憶されるテーブルを示す図である。

【図５】　本実施の形態に係る記憶ボリュームのペアを示す図である。

【図６】　本実施の形態に係る記憶デバイス制御装置において実行されるペア分割処理を説明するためのフローチャートである。

【図７】　本実施の形態に係る記憶デバイス制御装置において実行されるペア分割処理と入出力処理とを示すフローチャートである。

【図８】　第２の実施の形態に係る記憶デバイス制御装置における共有メモリに記憶されるテーブルを示す図である。

【図９】　第２の実施の形態に係る記憶デバイス制御装置において実行されるペア分割処理を説明するためのフローチャートである。

【符号の説明】

１００　情報処理装置
１１０　ＣＰＵ
１２０　メモリ
１２１　記憶デバイス管理プログラム
１２２　アプリケーションプログラム
１３０　ポート
１４０　記録媒体読取装置
１５０　入力装置
１６０　出力装置
１７０　記録媒体
２００　記憶デバイス制御装置
２１０　チャネルアダプタ
２１１　ＣＰＵ
２１２　キャッシュメモリ
２１３　制御メモリ
２１４　制御プログラム
２１５　ポート
２１６　バス
２２０　キャッシュメモリ

２３０　共有メモリ

２３１　コンシステンシグループ管理テーブル

２３２　ペア管理テーブル

２４０　ディスクアダプタ

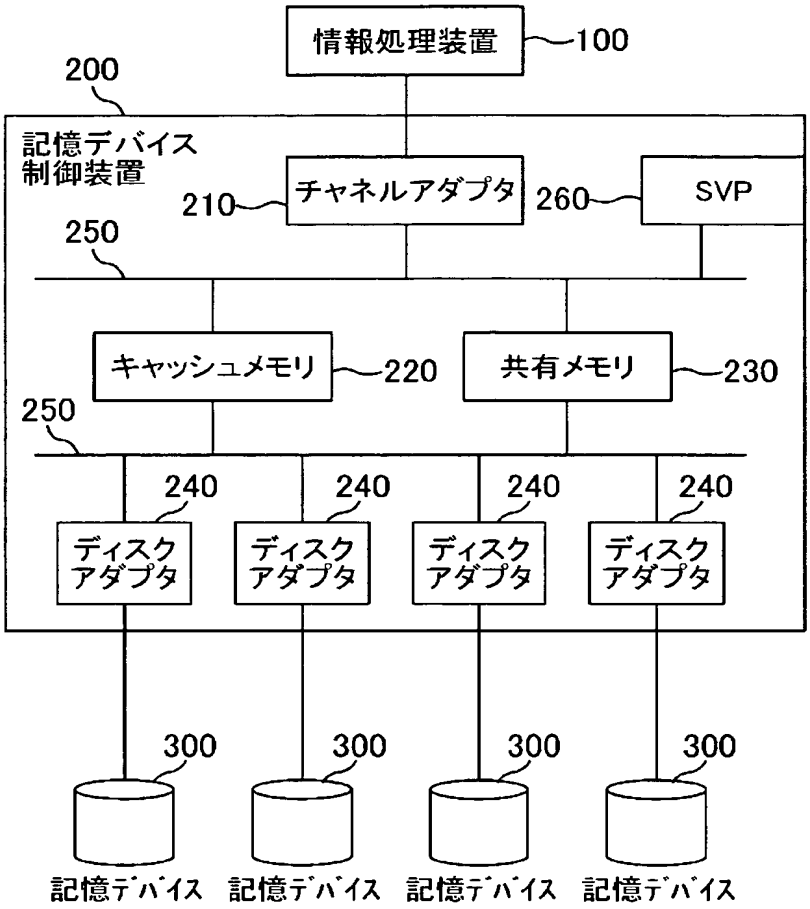
２５０　接続部

２６０　管理端末

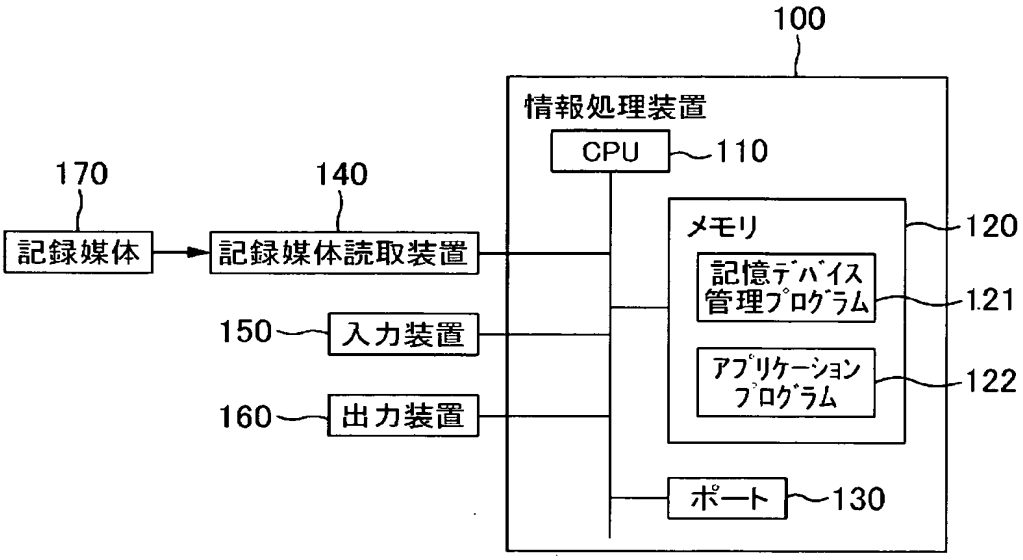
３００　記憶デバイス

【書類名】　図面

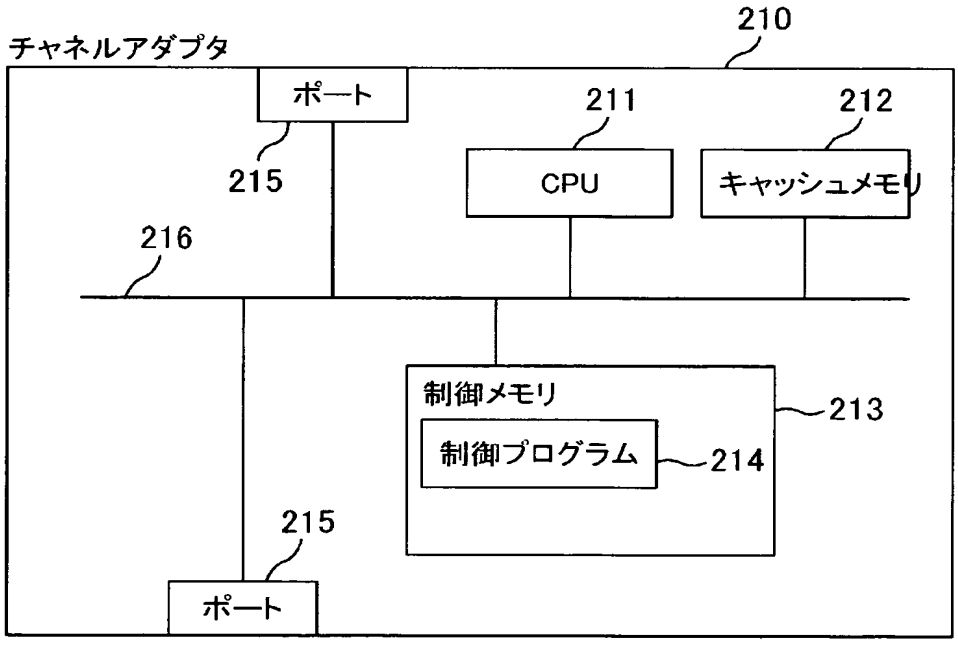
【図１】


情報処理装置
100
200
記憶デバイス
制御装置
210
チャネルアダプタ
260
SVP
250
キャッシュメモリ
220
共有メモリ
230
250
240
240
240
240
ディスク
アダプタ
ディスク
アダプタ
ディスク
アダプタ
ディスク
アダプタ
300
300
300
300
記憶デバイス
記憶デバイス
記憶デバイス
記憶デバイス

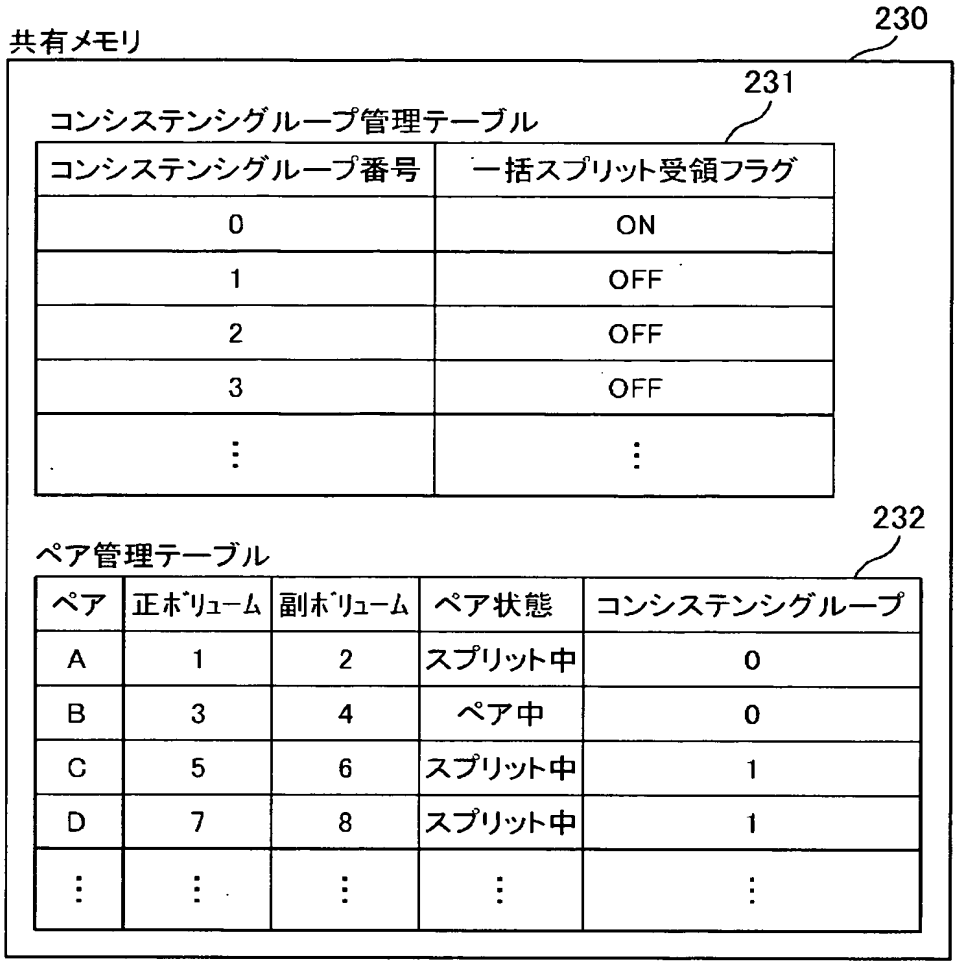
【図２】


100
情報処理装置
CPU
110
170
140
記録媒体
記録媒体読取装置
メモリ
120
記憶ﾃﾞﾊﾞｲｽ
管理ﾌﾟﾛｸﾞﾗﾑ
121
150
入力装置
ｱﾌﾟﾘｹｰｼｮﾝ
ﾌﾟﾛｸﾞﾗﾑ
122
160
出力装置
ポート
130

【図３】


210
チャネルアダプタ
ポート
211
212
215
CPU
キャッシュメモリ
216
制御メモリ
213
制御プログラム
214
215
ポート

【図４】



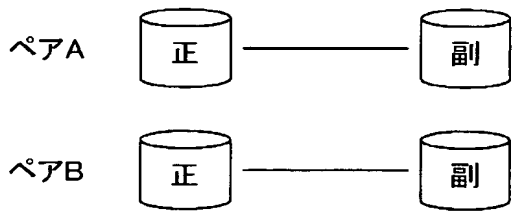
コンシステンシグループ管理テーブル

コンシステンシグループ番号	一括スプリット受領フラグ
0	ON
1	OFF
2	OFF
3	OFF
⋮	⋮

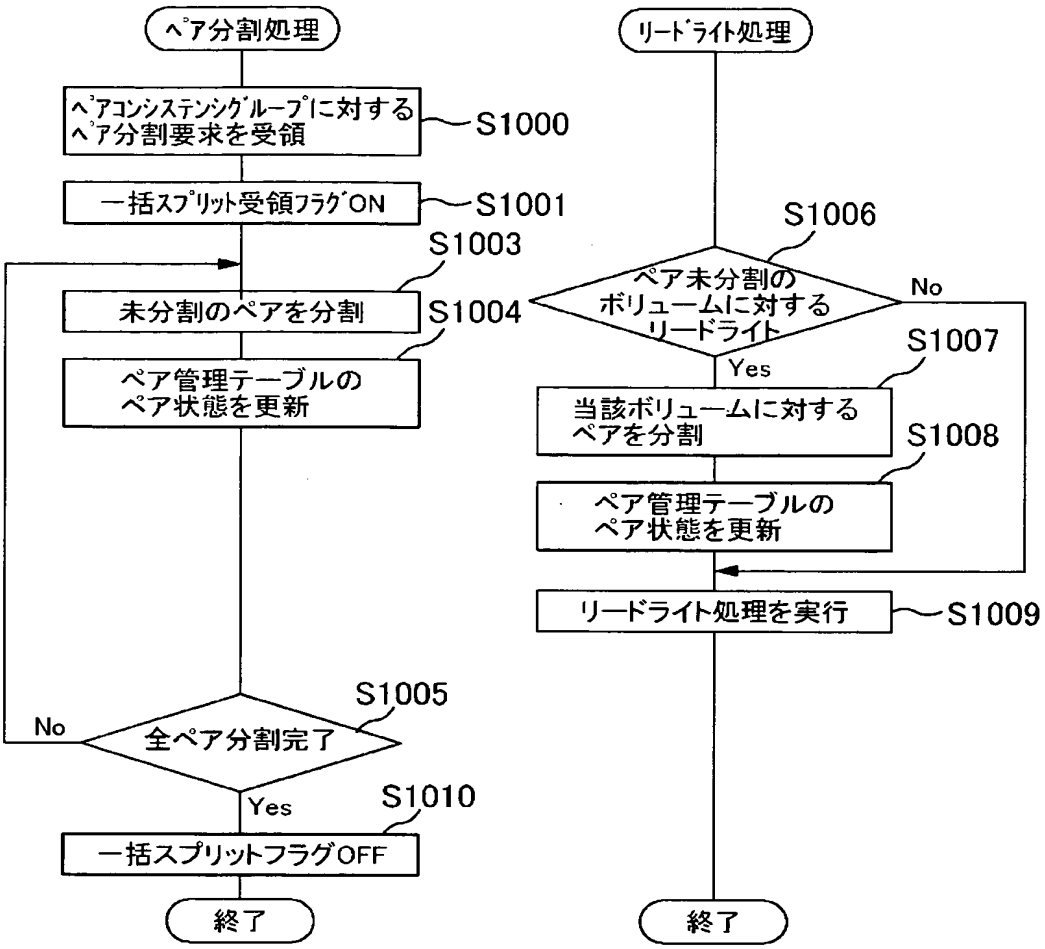
ペア管理テーブル

ペア	正ボリューム	副ボリューム	ペア状態	コンシステンシグループ
A	1	2	スプリット中	0
B	3	4	ペア中	0
C	5	6	スプリット中	1
D	7	8	スプリット中	1
⋮	⋮	⋮	⋮	⋮

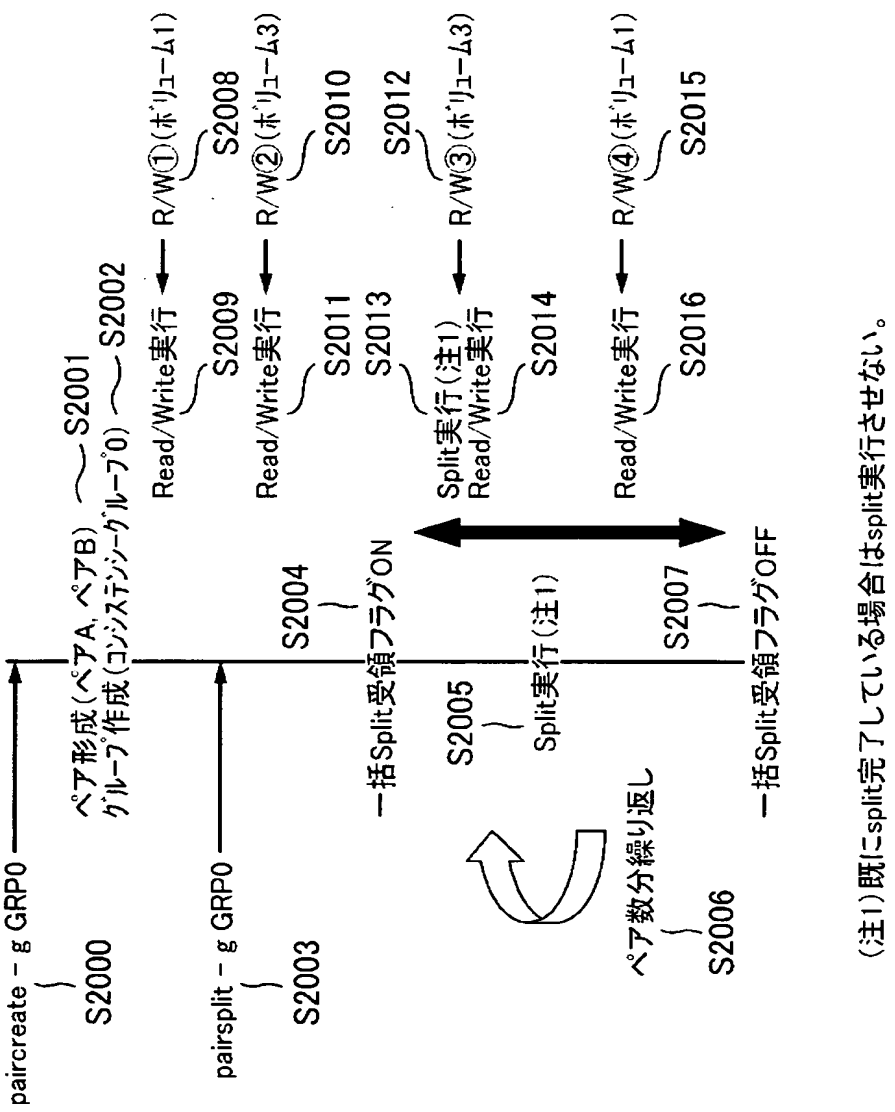
【図５】



【図６】

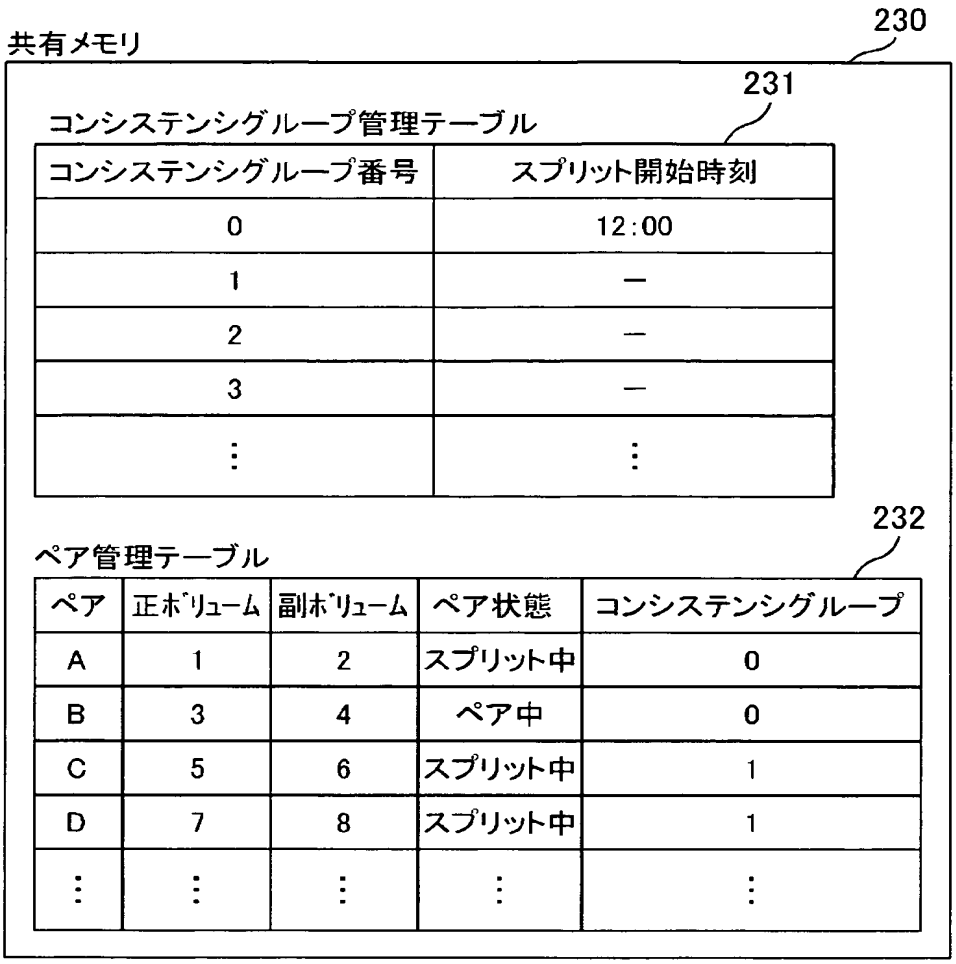

ペア分割処理
ペアコンシステンシグループに対するペア分割要求を受領
S1000
一括スプリット受領フラグON
S1001
S1003
未分割のペアを分割
S1004
ペア管理テーブルのペア状態を更新
S1005
No
全ペア分割完了
Yes
S1010
一括スプリットフラグOFF
終了
リードライト処理
S1006
ペア未分割のボリュームに対するリードライト
No
Yes
S1007
当該ボリュームに対するペアを分割
S1008
ペア管理テーブルのペア状態を更新
リードライト処理を実行
S1009
終了

【図７】



（注1）既にsplit完了している場合はsplit実行させない。

【図８】



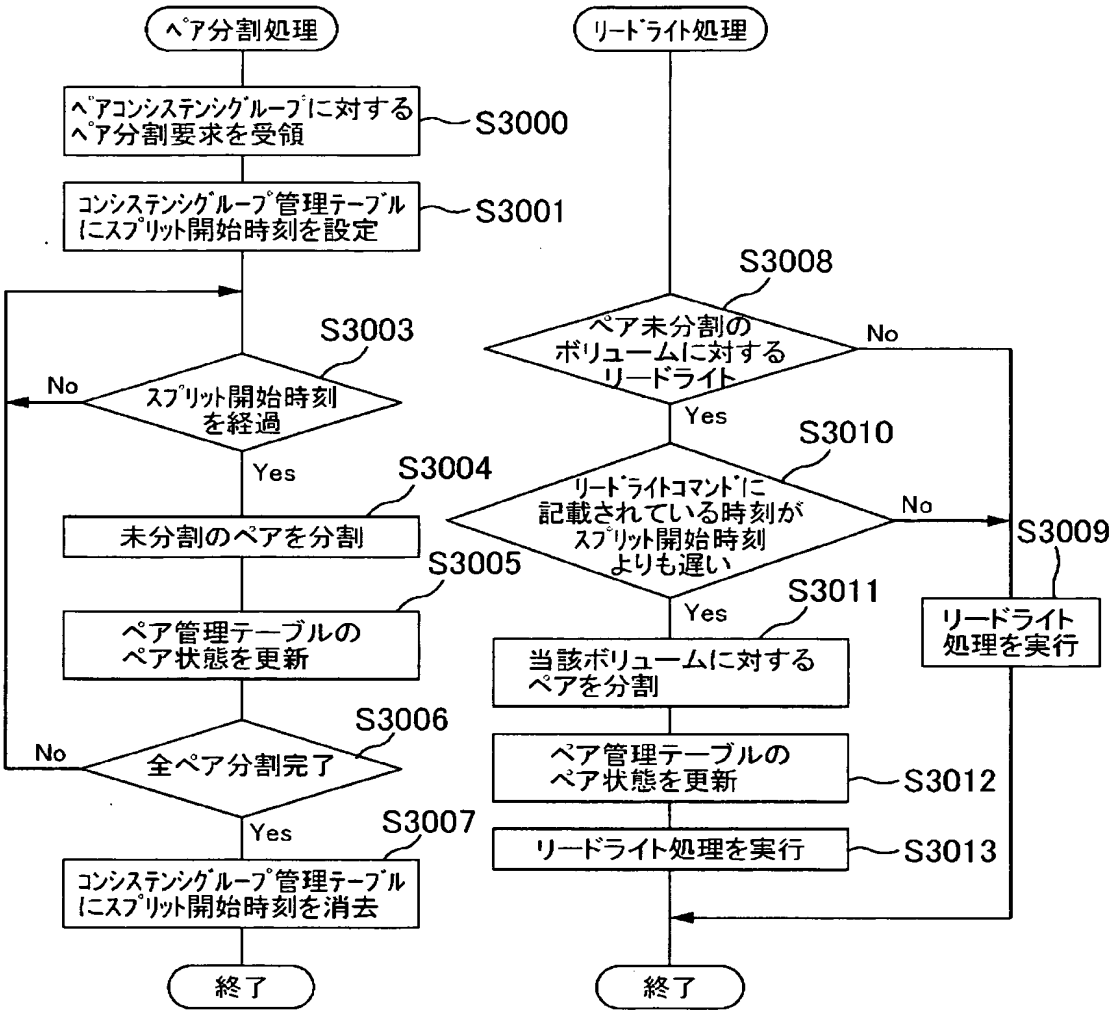
コンシステンシグループ管理テーブル

コンシステンシグループ番号	スプリット開始時刻
0	12:00
1	—
2	—
3	—
⋮	⋮

ペア管理テーブル

ペア	正ボリューム	副ボリューム	ペア状態	コンシステンシグループ
A	1	2	スプリット中	0
B	3	4	ペア中	0
C	5	6	スプリット中	1
D	7	8	スプリット中	1
⋮	⋮	⋮	⋮	⋮

【図９】


ペア分割処理
ペアコンシステンシグループに対するペア分割要求を受領
S3000
コンシステンシグループ管理テーブルにスプリット開始時刻を設定
S3001
S3003
スプリット開始時刻を経過
No
Yes
S3004
未分割のペアを分割
S3005
ペア管理テーブルのペア状態を更新
S3006
全ペア分割完了
No
Yes
S3007
コンシステンシグループ管理テーブルにスプリット開始時刻を消去
終了
リードライト処理
S3008
ペア未分割のボリュームに対するリードライト
No
Yes
S3010
リードライトコマンドに記載されている時刻がスプリット開始時刻よりも遅い
No
Yes
S3009
リードライト処理を実行
S3011
当該ボリュームに対するペアを分割
ペア管理テーブルのペア状態を更新
S3012
リードライト処理を実行
S3013
終了

【書類名】　　要約書

【要約】

【解決手段】　データを記憶するための記憶ボリュームを複数備える記憶デバイスとデータの入出力要求を行う情報処理装置とに接続され、情報処理装置から入出力要求を受信し記憶ボリュームに対するデータの入出力処理を行う記憶デバイス制御装置の制御方法であって、ある記憶ボリュームと、ある記憶ボリュームにデータが書き込まれると当該データの複製が書き込まれる他の記憶ボリュームとを対応付け、ある記憶ボリュームと他の記憶ボリュームとで構成されるペアを複数含んで構成されるペアグループを形成し、ペアグループを構成する各ペアの対応付けの解除を開始し、情報処理装置からの入出力要求が対応付けの解除を開始した後に発行されたものであるかどうかを判定し、対応付けの解除を開始した後に発行されたものである場合には対応付けの解除を行ってから入出力処理を行う。

【選択図】　　図１

特願２００２－３６６３７４

出　願　人　履　歴　情　報

識別番号　　［０００００５１０８］

１．変更年月日　　１９９０年　８月３１日
［変更理由］　　新規登録
住　所　　東京都千代田区神田駿河台４丁目６番地
氏　名　　株式会社日立製作所